Implementing data science to enhance water distribution system modelling. A case study of building hydraulic model of high-pressure zone in Wrocław

Wykorzystanie data science do usprawnienia modelowania hydraulicznego sieci wodociągowych. Przypadek budowy modelu strefy wysokiego ciśnienia we Wrocławiu

KAROL DYKIERT, WOJCIECH CIEŻAK, PIOTR BRÓDKA

DOI 10.36119/15.2025.10.6

This study presents the development and calibration of a hydraulic model for a high-pressure zone (HPZ) within the water distribution network (WDN) of Wrocław. The work focuses on advanced data processing methods and calibration techniques to improve model accuracy, drawing upon established concepts from both environmental engineering and data science. Our results indicate a very good model fit, although with minor inconsistencies regarding pipe roughness calibration. The overall approach demonstrates how the integration of statistical, hydraulic, and optimisation techniques can lead to robust and reliable WDN modelling, supporting both academic and operational needs.

Keywords: water distribution system modelling, data science, water distribution network, evolutionary algorithm, heuristic approach, artificial intelligence

W pracy omówiono budowę i kalibrację modelu hydraulicznego Strefy Wysokiego Ciśnienia (SWC), będącej częścią sieci wodociągowej miasta Wrocław. Opisano wybrane aspekty metodyki z zakresu analizy i obróbki danych wejściowych (preprocessingu), poprawiające dokładność modelu, bazujące na różnych technikach wykorzystywanych w inżynierii środowiska oraz data science. Otrzymane wyniki wskazują na wysoką jakość kalibracji modelu hydraulicznego, jakkolwiek zaobserwowano drobne nieścisłości w kalibracji chropowatości wewnętrznych ścian przewodów wodociągowych. Opracowana metodyka pokazuje, iż dzięki połączeniu metod statystycznych, hydraulicznych oraz technik optymalizacyjnych otrzymany model hydrauliczny jest wiarygodny i odpowiedni zarówno do pracy naukowej jak i zastosowań praktycznych.

Słowa kluczowe: model sieci wodociągowej, data science, sieć wodociągowa, algorytm ewolucyjny, heurystyka, sztuczna inteligencja

Introduction

Water Distribution Network (WDN) modelling is a specialised field within environmental engineering and WDN management that gained prominence towards the end of the 20th century with the development and subsequent widespread adoption of a modelling software. This advancement enabled researchers and engineers to explore new theories while allowing commercial enterprises to further enhance modelling capabilities. The rapid progress in information, computation, and communication technologies has provided access to vast amounts of data. However, it has also introduced challenges in handling diverse data types, such as water meter readings [2] and topological data from Geographic Information System (GIS) software, as described in detail by Di Nardo et al. [3].

The aim of this article is to highlight specific aspects of hydraulic modelling that required extensive processing, achieved through established concepts in data science. While these individual methods may not be particularly novel, their integration into a structured heuristic framework could offer valuable insights for modellers, lecturers, and students alike.

Case study area

The study area selected for our research is a high-pressure zone (HPZ) within the water distribution network in Wrocław (Figure 1). It comprises five District Meter Areas (DMAs), with pressure maintained by a pumping station. Although there are multiple connections to other DMAs, these are closed in daily operations. The HPZ was

MEng Karol Dykiert https://orcid.org/ 0009-0003-6168-7352 – Municipal Water and Sewerage Company (MPWiK) in Wrocław, Poland, Faculty of Environmental Engineering, Wrocław University of Science and Technology, Poland, Faculty of Information and Communication Technology, Wrocław University of Science and Technology, Poland
DEng Wojciech Cieżak https://orcid.org/ 0000-0001-6210-8728 – Faculty of Environmental Engineering, Wrocław University of Science and Technology, Poland
Assoc. Prof. DEng Piotr Bródka https://orcid.org/ 0000-0002-6474-0089 – Faculty of Information and Communication Technology, Wrocław University of Science and Technology, Poland.

Adres do korespondencji/ Corresponding author: karol.dykiert@pwr.edu.pl

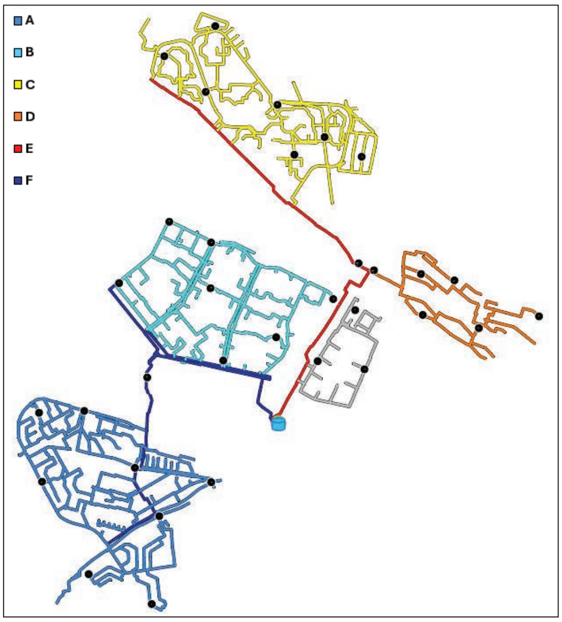

Figure 1.
Hydraulic model of HPZ in Wrocław. Pipe colours denote DMA assignments; grey pipes were excluded from the final calibration stage (see Results). Pressure measurement stations are marked with black dots. Flow was measured at DMA entry points and near the pumping station
Rys. 1. Model hydrauliczny SWC we Wrocławiu. Poszczególne strefy DMA zostały zaznaczone kolorami. Szara strefa nie została poddana ostatecznej kalibracji, co zostało wyjaśnione w Wynikach. Czarne punkty oznaczają lokalizację punktów pomiarowych

Table 1. Element counts in HPZ hydraulic model
Tabela 1. Kluczowe parametry obiektów modelu hydraulicznego SWC

Element type	Total count
Number of DMAs	5
Number of junctions	1247
Number of pipes	1341
Number of customers (water demand points)	1164
Number of water demand patterns	1034
Number of pumps	3
Number of reservoirs	1
Pipe length, km	68,2
Average total daily water demand (during 14-day period), m^3	8362
Number of temporary pressure measurement stations (TPMS)	30
TPMS sampling interval, minutes	1
Number of permanent pressure measurement stations (PPMS)	3
PPMS sampling interval, minutes	5

established to accommodate the significant number of high-rise apartment buildings that require elevated water pressure. The dataset provided by the Municipal Water and Sewerage Company in Wrocław (MPWiK) was supplemented by the temporary installation of hydrant pressure meters for the period of 11 September 2023 – 24 September 2023, covering a total of 14 days. Due to time constraints and various other factors, hydrant discharge tests were not conducted as part of the calibration process. Network conditions were assessed prior to the pressure measurements through operational hydrant tests, and subsequently by reviewing WDN malfunction and event logs. The model parameters and network characteristics are summarised in Table 1.

Materials and methods

The hydraulic model was developed using the commercial software MIKE+, provided by MPWiK. All external data preprocessing was conducted in Python, utilising the Pandas [13], [16], NumPy [8], and Matplotlib [10] libraries. The modelling process primarily followed the established methodology outlined by Walski [18]. A brief overview of the workflow is presented in Figure 2.

For the purpose of model calibration assessment, we employed the Root Mean Squared Error (RMSE) [9] and the coefficient of determination (R^2) [4]. RMSE quantifies the average deviation between observed and model-predicted values, while R^2 provides a statistical measure of how well the regression line approximates the actual data distribution. They are defined by the following equations:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}_i\right)^2}$$

Where n is the number of observations, y_i represents measured values, $\hat{y}_i$ represents values calculated by the model, $\bar{y}_i$ is the mean of observed values.

The key aspects of our approach are outlined below.

Initial pipe roughness assessment

The initial task involved assigning preliminary pipe roughness values. In accordance with the established textbook, we employed research from 1981 [12] that detailed Hazen-Williams C-factor values for various materials, with particular emphasis on cast iron pipes of differing ages and subjected to varying water quality levels, characterised by their degree of attack (corrosiveness). This consideration was especially critical, given that approximately 45% of the pipes in the area are composed of cast iron. At the same time, the remaining materials that exhibit scaling are present in considerably lower proportions. Based on this research, we developed an initial framework for the assignment of pipe roughness. We interpolated roughness values corresponding to moderate attack for cast iron pipes across different diameters to establish a range of assignable values (Figure 3).

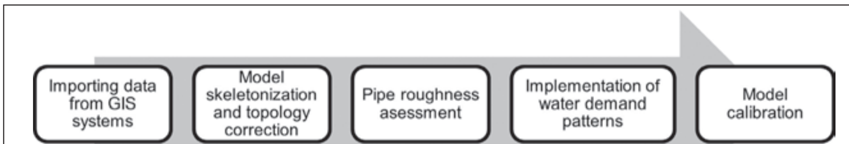

Figure 2.
A summary of WDN model development process
Rys. 2. Uproszczony schemat budowy modelu hydraulicznego sieci wodociągowej

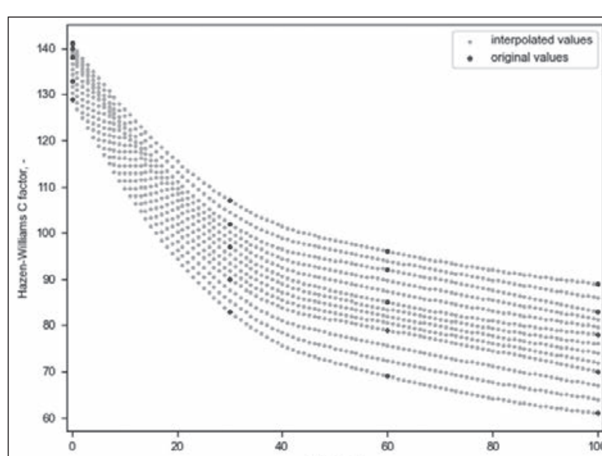

Figure 3.
An example of C-factor calculated for cast iron pipes of various diameters
Rys. 3. Przykład wartości współczynnika C Hazena-Williamsa obliczonego dla przewodów wodociągowych o różnych średnicach

The roughness values for other materials, including ductile iron, steel, and asbestos cement, were subsequently derived. Owing to the scarcity of data on the effects of age and water quality for these materials, we elected to apply the curves developed for cast iron as an approximation. This decision was deemed acceptable given the relatively low prevalence of other materials in the area, rendering them a lower priority.

Although the choice of the Hazen–Williams C-factor may appear unconventional, given its predominant application in the United States, the referenced study [12] also employed this parameter and provided valuable insights for the initial assessment of pipe roughness and the parameterisation of the roughness calibration algorithm. While the C-factor is limited to turbulent flow regimes, this limitation was considered acceptable, as pipe roughness is primarily of concern under such conditions.

Developing water demand patterns

The development of water demand patterns required particular attention. As the model was being constructed, we aimed to fully utilise the available data, prioritising volumetric and individual demand patterns to accurately replicate water flow at the time of our measurements. MPWiK had installed water meters with more frequent reading intervals, enabling hourly measurements. However, not all water meters had been upgraded at the time of the pressure measurements. To expand the dataset, we retrieved readings from June 2024, as a quantitative analysis demonstrated a strong similarity in water consumption during this period. Despite this approach, a considerable number of missing values remained for individual water meters. Following discussion, we decided to fill missing values for time series where gaps did not exceed two days within a 14-day period. To achieve this, we employed the concept closest fit [6], a method that segments the dataset into predefined classes and imputes missing values using the closest available record within the same class. In our case, the dataset was divided into 72 classes, obtained by assigning one of three labels – Workday, Saturday, or Sunday – to each full day and further subdividing these categories by the hour of the day. This approach ensured that both diurnal and hourly variations in water demand were preserved.

The remaining nodes with present water demand received a general pattern based on total water consumption measured at the pumping station. The constant average hourly consumption was multiplied by a percentage value in each timestep.

Reference data formatting and resampling

As previously mentioned, the development of the hydraulic model was centred on the additional pressure measurements. After integrating these data with records from permanent pressure and flow monitoring points, we constructed a dataset comprising time series with varying timesteps. This variation posed challenges for comparison and model evaluation, as higher-frequency data exhibited greater noise and a higher number of outliers, affecting quality metrics and making visual representation less interpretable. To address this issue, we applied a smoothing method using bin means [7], where each bin represented a full hour within the 14-day measurement period. Smoothing by bin means is similar to the concept closest fit; however, instead of filling in missing values, it replaces all records within a class with a single value. In this case, the replacement value is the mean calculated from all data points within the respective class. This adjustment was necessary owing to the difference in timestep lengths between the water meters and pressure meters—one hour and one minute, respectively. Consequently, a one-hour timestep was adopted for the model.

For the purpose of calibration, we restricted the dataset to periods when the pumping station was fully operational, excluding nighttime when the station was shut down. This approach helped eliminate potential biases that could significantly impact the calibration process. Retaining data characterised by substantially lower pressure and flow rates could mislead the autocalibration algorithm and compromise its overall performance. Moreover, the Hazen–Williams C-factor applies exclusively to turbulent flow regimes and is therefore unsuitable for the very low flow rates observed at night. Figure 4. presents a comparison between the original raw data and the processed data used for calibration, illustrated as a Hydraulic Grade Line (HGL) for one of the temporary pressure measurement stations.

Pipe roughness calibration

The final aspect requiring extensive work was pipe roughness calibration. Unlike other parameters, pipe roughness presents a significant challenge due to the highly irregular and unpredictable nature of pipe scaling. Consequently, the calibration of pipe roughness remains an active area of scientific research, primarily focused on developing and implementing various methods to address this issue.

For our study, we adopted a methodology inspired by the work of Zhao et. al. [20], with modifications tailored to our

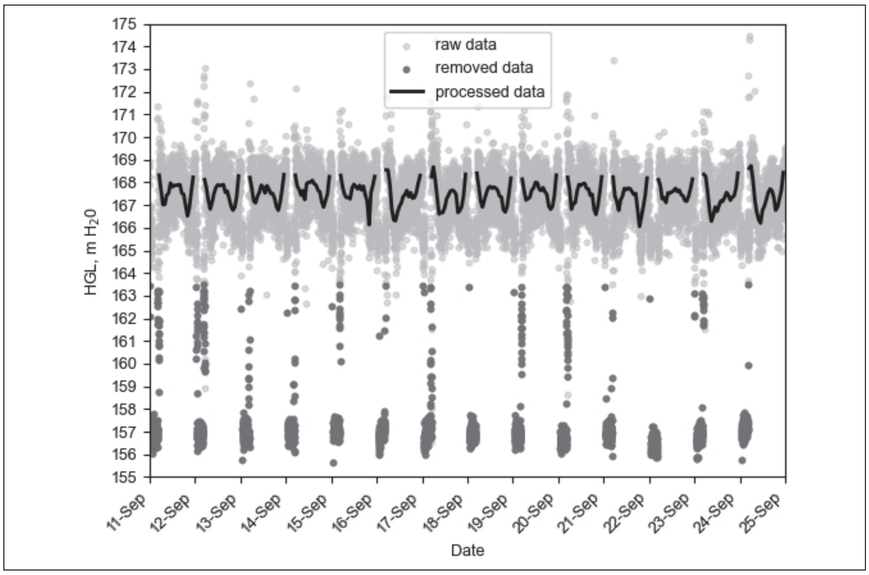

Figure 4.
A comparison between the original raw data and the processed data used for calibration, for one of the measurement points. Notably, the visible drops in water pressure – resulting from the shutdown of the pumping station – have been removed. For clarity, the original timestamps have been retained in the processed dataset

Rys. 4. Porównanie surowych danych z danymi przygotowanymi do kalibracji, dla jednego z punktów pomiarowych. Zauważalny spadek ciśnienia, wynikający z wyłączenia przepompowni, został usunięty. Dla klarowności wykresu oryginalne stemple czasowe zostały zachowane w przeliczonym szeregu czasowym

specific case. We selected the Shuffled Complex Evolution (SCE) algorithm [5], given well-established success of evolutionary algorithms in optimisation techniques for water resource management [11], including WDN model calibration [1],[14], [17],[19], as well as its availability within the provided modelling software.

SCE is a global optimisation method designed for continuous optimisation problems and is widely applied in hydrological modelling and environmental engineering. The algorithm operates by evolving multiple populations (complexes) of potential solutions through iterative shuffling and evolution steps. In the modelling software, model fitness is assessed by computing an aggregated objective function – a weighted sum of objectives specified by the modeller.

Building on previous research [20] and network analysis, we developed a methodology for parameterising SCE. The algorithm's standard parameters were implemented according to recommendations from the original research [5]. Decision variables were defined based on two criteria: pipe diameter (D) and DMA, resulting in a total of 27 decision variables. Each decision variable was assigned a range of acceptable pipe roughness values, based on information from Lamont's work [12]. The target objectives were established using reference data, with weighting factors adjusted through preliminary data analysis. In this process, we reduced the weights of objectives affected by a higher number of outliers and anomalies.

Results and discussion

The successful development and calibration of the HPZ hydraulic model were achieved through the application of the aforementioned methods. The use of the concept of closest fit allowed nearly 88% of all water demand points (1024 out of 1164 water demand patterns) to receive individual water demand patterns, enabling an accurate recreation of water flow during the selected calibration period. Figure 5. shows a comparison between the flow measured at the pumping station and the calculated flow.

Processing the reference data provided deeper insights into established patterns, allowing for the refinement of various parameters, such as the weights assigned to decision variables. Regarding model fitness, the RMSE values calculated for all reference water pressure data ranged from 0,29 to 1,3 m H_2O, while the R^2 varied between 0,92 and 0,99. The median RMSE across all measurement points was 0,46, while the median R^2 value was 0,98, further reinforcing the overall accuracy of the model. Figure 6. shows the comparison between calculated and measured values for one of the measurement points.

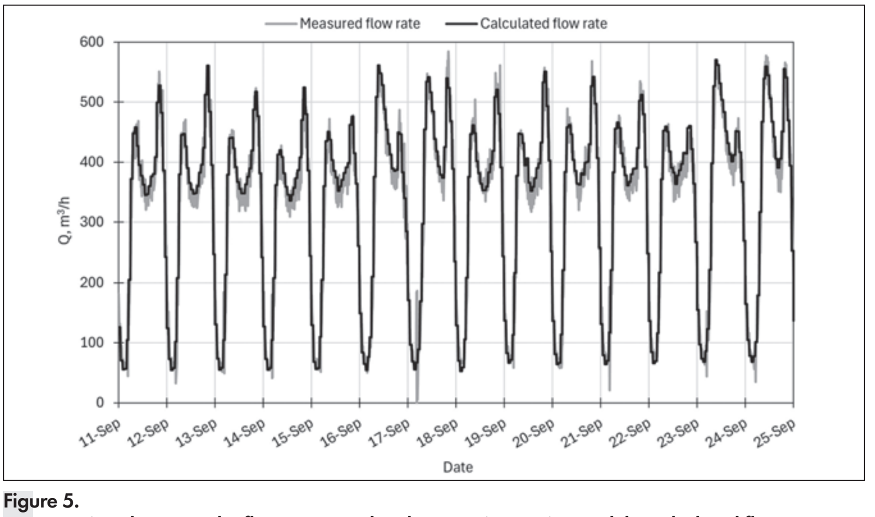

Figure 5.
A comparison between the flow measured at the pumping station and the calculated flow
Rys. 5. Porównanie zmierzonego natężenia przepływu na przepompowni z przepływem obliczonym w modelu

These results indicate a very good model fit, with the majority of the dataset exhibiting minimal deviation from observed values. Several factors contributed to discrepancies between the model and the reference data, including:

- Differences in network depth between the model and the actual system, resulting from the use of a Digital Elevation Model as a data source for node elevation,
- A pipe burst causing a significant pressure drop at one of the temporary pressure measurement stations,
- Other minor anomalies that were difficult to accurately implement in the model.

The complete list of decision variables used for pipe roughness calibration is presented in Table 2.

One of the key elements requiring clarification is the notable variation in the range of C-factors assigned to decision variables in locations E and F compared to other zones. Preliminary calibration results indicated that the initially assigned C-factors in these locations (defined in Step 1.) were nearly adequate, which led us to restrict the variation range for the optimisation process. For the same reason, one of the DMAs was excluded from calibration altogether. Furthermore, pipes with diameters smaller than 100 mm were omitted due to their marginal share in the HPZ. Analysis of the calibration outcomes revealed several important trends:

- All 300 mm pipes were assigned a C-factor of 43, the minimum value permitted by the algorithm.

Figure 6.
Correspondence between calculated and measured values, exemplified at one of the measurement points. Part a) presents the data used for calibration, excluding periods when the pumping station was shut down at night, while part b) illustrates the full two-week simulation
Rys. 6. Korelacja między wartościami obliczonymi a zmierzonymi w jednym z punktów pomiarowych. Część a) przedstawia dane wykorzystane do kalibracji, z wyłączeniem okresów, w których stacja pomp była wyłączona w nocy, natomiast część b) przedstawia pełne dwa tygodnie symulacji

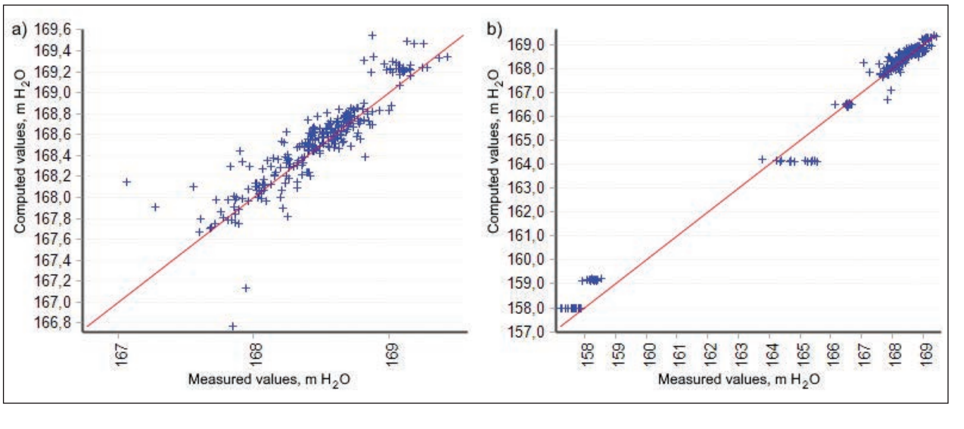

Table 2. List of decision variables defined for pipe roughness calibration. Each variable is characterised by pipe diameter and location, the latter determined based on DMA boundaries. Transmission mains were further subdivided into two separate topological regions for calibration purposes. The table includes calibrated values of the Hazen–Williams C-factor, along with the minimum and maximum bounds assigned to each variable, which defined the search space for the SCE algorithm
Tabela 2. Lista zmiennych decyzyjnych zdefiniowanych dla algorytmu kalibrującego chropowatość. Każda zmienna została określona bazując na średnicy oraz lokalizacji, wyznaczonej przez granice DMA. Sieci magistralne zostały podzielona na dwie oddzielne lokacje. Tabela zawiera dobrane wartości współczynnika C Hazena-Williamsa dla każdej zmiennej wraz ze zdefiniowanym przedziałem wartości, z którego algorytm SCE dobierał wartość podczas kalibracji

Decision variable no.	Diameter (D), mm	Location	C minimum, –	C maximum, –	C calibrated, –
1	100	A	28	76	36,9
2	125	A	31	78	40,9
3	150	A	34	80	34,0
4	200	A	38	83	83,0
5	250	A	41	85	78,5
6	300	A	43	87	43,0
7	100	B	28	76	53,7
8	150	B	34	80	39,7
9	200	B	38	83	60,1
10	300	B	43	87	43,0
11	400	B	46	93	61,5
12	100	C	28	76	28,0
13	150	C	34	80	34,0
14	200	C	38	83	38,2
15	300	C	43	87	43,0
16	100	D	28	76	48,4
17	150	D	34	80	36,9
18	200	D	38	83	38,0
19	300	D	43	87	43,0
21	500	E	92	109	92,0
22	600	E	94	111	95,6
23	800	E	96	113	102,2
20	1000	E	108	117	117,0
24	400	F	90	108	104,2
25	500	F	92	108	92,0
26	600	F	94	111	95,0
27	800	F	96	113	99,3

- There was no clear correlation between pipe diameter and calibrated C-factor values. In some cases, smaller-diameter pipes were assigned higher C-factors than larger ones, a counterintuitive result highlighted in Figure 7.
- Out of 27 decision variables, 10 were assigned the minimum value and 2 the maximum, suggesting the algorithm gravitated towards boundary solutions for a significant portion of the variables.

To facilitate further analysis of the relationship between roughness and pipe diameter, the calibrated C-factors were converted into absolute roughness (ε) and subsequently into relative roughness (ε/D), allowing for a normalised comparison. This transformation was applied at the link level within the model, rather than at the level of decision variables, ensuring a finer granularity of analysis. The outcomes of this process are depicted in Figure 8., which presents boxplots generated for each pipe diameter.

The calibrated relative roughness values exhibit a clear dependence on pipe diameter. Notably, two distinct trends can be observed: pipes with a diameter of 300 mm and smaller exhibit significantly higher relative roughness values compared to larger pipes. This observation led us to conclude that the initial pipe roughness assessment was sufficiently accurate for larger pipes, which were therefore excluded from the final roughness calibration using the SCE. The better condition of these larger-diameter transmission mains may be attributed to more frequent renovations

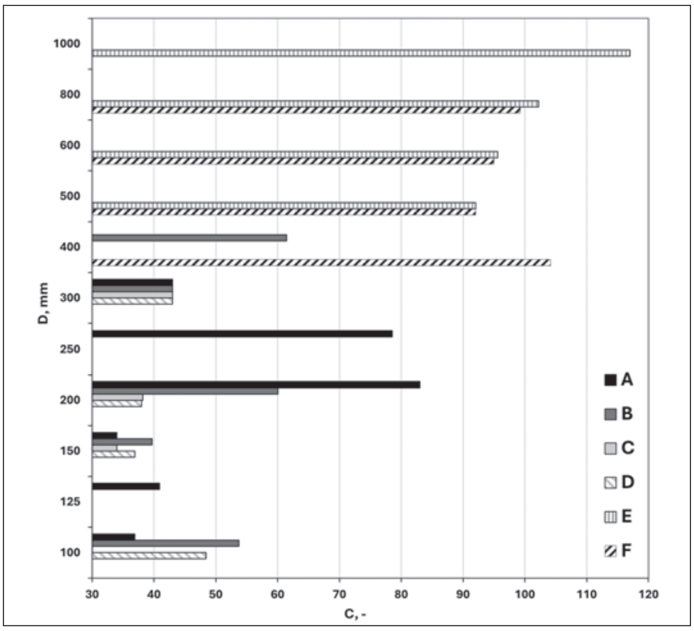

Figure 7.
Values of calibrated C-factor for all decision variables
Rys. 7. Dobrane wartości współczynnika C dla wszystkich zmiennych decyzyjnych

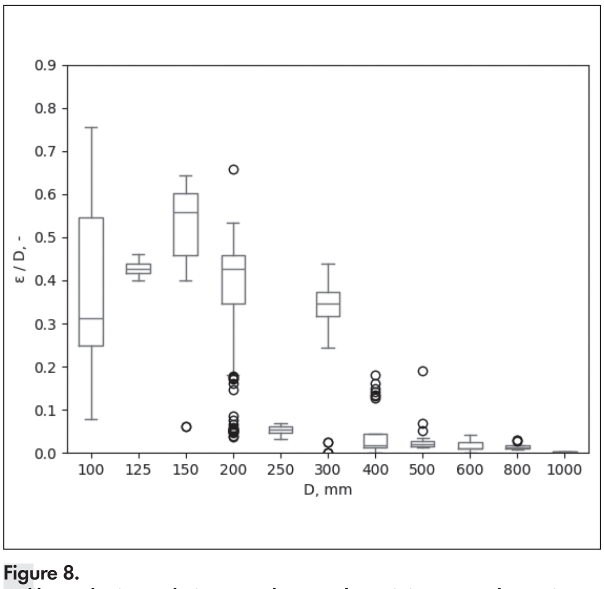

Figure 8.
Calibrated pipe relative roughness. The minimum and maximum values were identified within 1.5 times the interquartile range
Rys. 8. Chropowatość względna skalibrowanej sieci wodociągowej. Minimum i maksimum zostało określone na podstawie 1,5-krotnej wartości rozstępu międzykwartylowego

and replacements, for which some supporting documentation was available. However, this reasoning does not appear to apply to smaller-diameter pipes, which function primarily as distribution mains.

Furthermore, while relative roughness is generally expected to decrease with increasing pipe diameter, the distribution of values for certain diameters appears somewhat irregular, particularly for 100 mm and 200 mm. This variability may be linked to the specific locations of these pipes within the network, specific placement of pressure sensors and the network's looped topology, which poses challenges for accurate calibration. But it may also indicate that the calibration algorithm prioritised adjustments for 300 mm pipes while making minimal modifications to the rest.

Conclusions

Our methodology integrates a range of data processing and analysis techniques to ensure the accuracy and reliability of the hydraulic model. To maximise the volume of usable data for calibration, we employed methods such as concept closest fit and bin smoothing. These approaches significantly enhanced data completeness, contributing to the overall strong fit of the model.

With regard to the SCE algorithm, our analysis revealed that it predominantly adjusted roughness values for larger-diameter pipes, while the calibration of smaller pipes remained inconsistent. This discrepancy suggests that the algorithm may have prioritised certain pipe categories, potentially overlooking finer-scale variations in roughness across the network.

In the final stage of calibration, pipe material was not explicitly considered, as its impact was deemed low priority due to the relatively small proportion of aging materials other than cast iron within the HPZ. While incorporating material-specific roughness adjustments could have further improved calibration accuracy, it is also necessary to examine additional factors, including both hydraulic and topological parameters.

While our methodology simplified the data and produced a good model fit, it inevitably sacrificed some nuances for the sake of clarity. Therefore, the calibration should be validated with additional pressure and flow measurements, potentially including hydrant discharge tests where feasible. Further analysis of flow rates could help identify areas of the network that would benefit from such tests. In addition, the chosen one-hour timestep may not be optimal, given the dynamic nature of water pressure, flow, and demand. Future work could consider more detailed measurements to refine and improve our approach.

Addressing these challenges will require further research, particularly in developing a more comprehensive mathematical framework for defining decision variables in roughness calibration. Future studies will explore refined methodologies that account for a broader set of influencing factors, aiming to enhance the precision and robustness of the calibration process.

Acknowledgments

The modelling software and relevant data were provided by MPWiK. In the preparation of this article, we utilised ChatGPT [15] for grammar correction and vocabulary refinement.

This methodology forms part of the implementation doctorate *Optimisation of Water Network Pressure Using Artificial Intelligence* and may be subject to further modifications. The project received additional funding from the Polish Ministry of Science and Higher Education under the *Implementation Doctorate* Programme, 6th edition.

Hydraulic model development, calibration, and the visualisation shown in Figure 6. were carried out using Mike+ 2024, developed by DHI.

REFERENCES

[1] Bylka J, Bałut A, Brodziak R, Zakrzewski P. Nodal Demand Control Genetic Algorithm for Water Supply Systems. *Instal.* 2022;10. doi:10.36119/15.2022.10.8
[2] Cichoń T, Królikowska J. The role and importance of smart water meters in the water supply system [in Polish]. *Rola i znaczenie inteligentnych wodomierzy w systemie zaopatrzenia w wodę. Instal.* 2024;12. doi:10.36119/15.2024.12.7
[3] Di Nardo A, Boccelli DL, Herrera M, et al. Smart Urban Water Networks: Solutions, Trends and challenges. *Water.* 2021;13(4):501. doi:10.3390/w13040501
[4] Dodge Y. *The Concise Encyclopedia of Statistics.*; 2008. doi:10.1007/978-0-387-32833-1
[5] Duan QY, Gupta VK, Sorooshian S. Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications.* 1993;76(3):501-521. doi:10.1007/bf00939380
[6] Grzymala-Busse JW, Grzymala-Busse WJ. Handling missing attribute values. In: *Springer eBooks.* ; 2009:33-51. doi:10.1007/978-0-387-09823-4_3
[7] Han J, Kamber M, Pei J. Data preprocessing. In: *Data Mining Concepts and Techniques.* Elsevier eBooks; 2012:83-124. doi:10.1016/b978-0-12-381479-1.00003-4
[8] Harris CR, Millman KJ, Van Der Walt SJ, et al. Array programming with NumPy. *Nature.* 2020;585(7825):357-362. doi:10.1038/s41586-020-2649-2
[9] Hunter JD. MatPlotLib: a 2D Graphics environment. *Computing in Science & Engineering.* 2007;9(3):90-95. doi:10.1109/mcse.2007.55
[10] Hodson TO. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development.* 2022;15(14):5481-5487. doi:10.5194/gmd-15-5481-2022
[11] Kumar V, Yadav SM. A state-of-the-Art review of heuristic and metaheuristic optimization techniques for the management of water resources. *Water Science & Technology Water Supply.* 2022;22(4):3702-3728. doi:10.2166/ws.2022.010
[12] Lamont PA. Common pipe flow formulas compared with the theory of roughness. *American Water Works Association.* 1981;73(5):274-280. doi:10.1002/j.1551-8833.1981.tb04704.x
[13] McKinney W. Data structures for statistical computing in Python. *Proceedings of the Python in Science Conferences.* Published online January 1, 2010:56-61. doi:10.25080/majora-92bf1922-00a
[14] Nicolini M, Falcomer L. *Genetic Algorithm for Calibration and Leakage Identification in Water Distribution System.* 2020 3rd IEEE International Conference on Knowledge Innovation and Invention (ICKII); 2020:273-276. doi:10.1109/ickii50300.2020.9318899
[15] OpenAI. ChatGPT (GPT-4o). Accessed April 7, 2025. https://chat.openai.com/chat.
[16] Pandas Development Team. pandas-dev/pandas: Pandas. *Zenodo.* Published online February 2020. doi:10.5281/zenodo.15831829
[17] Shiu CC, Chung CC, Chiang T. Enhancing the EPANET Hydraulic Model through Genetic Algorithm Optimization of Pipe Roughness Coefficients. *Water Resources Management.* 2023;38(1):323-341. doi:10.1007/s11269-023-03672-0
[18] Walski TM, Chase DV, Savic DA, Grayman W, Beckwith S, Koelle E. *Advanced water Distribution Modeling and Management.*; 2003. https://ecommons.udayton.edu/cee_fac_pub/18/
[19] Wu ZY, Clark C. Evolving effective hydraulic model for municipal water systems. *Water Resources Management.* 2008;23(1):117-136. doi:10.1007/s11269-008-9267-3
[20] Zhao Q, Wu W, Simpson AR, Willis A. Simpler is Better—Calibration of pipe roughness in water distribution systems. *Water.* 2022;14(20):3276. doi:10.3390/w14203276